

Bayesian inferential framework for diagnosis of non-stationary systems

Vadim N. Smelyanskiy^a, Dmitry G. Luchinsky^a, Andrea Duggento^b and Peter V. E. McClintock^b

^aNASA Ames Research Center, MS 269-3, Moffett Field, CA 94035-1000, U.S.

^bPhysics Department, Lancaster University, LA1 4YB, Lancaster, U.K.

ABSTRACT

A Bayesian framework for parameter inference in non-stationary, nonlinear, stochastic, dynamical systems is introduced. It is applied to decode time variation of control parameters from time-series data modelling physiological signals. In this context a system of FitzHugh-Nagumo (FHN) oscillators is considered, for which synthetically generated signals are mixed via a measurement matrix. For each oscillator only one of the dynamical variables is assumed to be measured, while another variable remains hidden (unobservable). The control parameter for each FHN oscillator is varying in time. It is shown that the proposed approach allows one: (i) to reconstruct both unmeasured (hidden) variables of the FHN oscillators and the model parameters, (ii) to detect stepwise changes of control parameters for each oscillator, and (iii) to follow a continuous evolution of the control parameters in the quasi-adiabatic limit.

Keywords: Nonlinear time-series analysis, Bayesian inference, varying parameters, FitzHugh-Nagumo, measurement matrix.

1. INTRODUCTION

The ubiquity of noisy nonlinear systems in nature has led to the use of stochastic nonlinear dynamical models for observed phenomena across many scientific disciplines, including the modelling of biological nonstationary signals. Important “hidden” features of a model such as coupling coefficients between the dynamical degrees of freedom can be very difficult to extract due to the presence of intrinsic dynamical noise and the intricate interplay between noise and nonlinearity.

The problem becomes even more challenging if we want to detect and characterize the nonstationarity of the system following the evolution of system’s parameters. Fast (almost real-time) detection of parameters’ evolution is essential in the identification, diagnostic, and prognostic of the time evolution of complex dynamical systems. This is especially true when dealing with biological signals.

In our earlier work¹ an efficient technique of Bayesian inference of nonlinear noise-driven dynamical models that guarantees optimum compensation of dynamical noise-induced errors for continuous systems and avoids extensive numerical optimization was introduced. In the present paper we extend earlier results and introduce a Bayesian framework for inferring parameters of continuous non-stationary systems, and in particular for the long standing problem of reconstruction of parameters for neuronal systems.

Neuronal signals are a good example, for which the Bayesian inference is especially advantageous: often the system’s dynamical details are known only approximately. There is a strong influence from internal noise but, most of all, the coefficients and coupling terms are in general not constant, and so a fast and “elastic in adaptation” technique is needed to detect changes of the system.

In this context a system of FitzHugh-Nagumo (FHN) oscillators is considered, for which synthetically generated signals are mixed via a measurement matrix. For each oscillator only one of the dynamical variables is assumed to be measured, while another variable remains hidden (unobservable). The control parameter for each FHN oscillator is varying in time. The goal is to decode a control parameter for each signal. Note that

Further author information: (Send correspondence to A.D.)

A.D.: E-mail: a.duggento@lancaster.ac.uk, Telephone: +44 (0) 1524 593206, Address: Department of Physics, Lancaster University, Lancaster LA1 4YB, U.K.

the system of FitzHugh-Nagumo oscillators is one of the most useful models²⁻⁴ in explaining a lot of different biological dynamics. It can present spontaneous oscillatory firing patterns⁵ and it can be used to characterize electrical signal propagation in nerve fibres⁶ and in cardiac tissue.⁷⁻⁹

The paper is organized as follows. First, we describe how the Bayesian algorithm is applied to a single oscillator, and then we discuss a multi-dimensional system, where every signal coming from a FitzHugh-Nagumo oscillator is linearly mixed with the other ones through a specific mixing measurement matrix. We then provide practical demonstration of how model parameters and noise coefficients can be inferred in this situation, and consider the advantages and limitation of this present technique.

2. BACKGROUND ON BAYESIAN INFERENCE

Let us suppose we have to analyse a signal coming from an L -dimensional system of the form

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}|\mathbf{c}) + \boldsymbol{\xi}(t) \\ \langle \boldsymbol{\xi}(t) \rangle &= 0, \quad \langle \boldsymbol{\xi}(t) \boldsymbol{\xi}^T(t') \rangle = \hat{\mathbf{D}} \delta(t - t')\end{aligned}\tag{1}$$

where \mathbf{c} is a vector of unknown parameters and $\boldsymbol{\xi}(t)$ is a white Gaussian additive dynamical noise characterized by diffusion matrix $\hat{\mathbf{D}}$. The task is to decode the optimal set of parameters $\mathcal{M} = \{\mathbf{c}, \hat{\mathbf{D}}\}$, given a set of measurement \mathcal{X} .

A priori knowledge of \mathcal{M} is summarized in the so-called *prior* probability density function (PDF) $p_{\text{pr}}(\mathcal{M})$. Then the time-series \mathcal{X} is acquired from the experimental set up, and the new gained information is used to compute the *a posteriori* PDF of the model parameters, namely $p_{\text{post}}(\mathcal{M}|\mathcal{X})$. The relationship between the two distribution $p_{\text{pr}}(\mathcal{M})$ and $p_{\text{post}}(\mathcal{M}|\mathcal{X})$ is given by the Bayes' theorem:

$$p_{\text{post}}(\mathcal{M}|\mathcal{X}) = \frac{\ell(\mathcal{X}|\mathcal{M}) p_{\text{pr}}(\mathcal{M})}{\int \ell(\mathcal{X}|\mathcal{M}) p_{\text{pr}}(\mathcal{M}) d\mathcal{M}}.\tag{2}$$

The PDF $\ell(\mathcal{X}|\mathcal{M})$ (also called *likelihood* of \mathcal{X}) is the conditional probability of the occurrence of measurements \mathcal{X} when the set of parameters \mathcal{M} is given.

The procedure can be applied iteratively for a sequence of data blocks $\{\mathcal{X}_1, \mathcal{X}_2, \dots\}$: the posterior for one block is the prior for the next block:

$$p_{\text{pr}}(\mathcal{M}_{i+1}) = p_{\text{post}}(\mathcal{M}_i|\mathcal{X}_i)$$

The densities p_{post} for each \mathcal{M}_i are a sequence of functions that become sharper and sharper functions that peak at the solution of the inference problem.

One of the main goals of dynamical inference is to find the correct analytical expression for the likelihood and to introduce an efficient algorithm for optimization of the posterior distribution with respect to the model parameters. Both problems are very non trivial in the context of nonlinear dynamical stochastic systems. In the earlier research the correct analytical form for $\ell(\mathcal{X}|\mathcal{M})$ could not be found and optimization was mainly relying on extensive numerical methods.

In a recent work¹ we have introduced (and verified in application to the dynamical inference of cardio-vascular system⁷) analytical solution for a wide class of non-linear stochastic systems which does not require extensive numerical calculations and provides an optimal compensation for the noise-induced errors. One of the keys of the novel approach is to write the expression for the likelihood in the form of a path integral over the random trajectories of the system:

$$\ell(\mathcal{X}|\mathcal{M}) = \int_{\mathbf{x}(t_i)}^{\mathbf{x}(t_f)} \mathcal{F}_{\mathcal{M}}[\mathbf{x}(t)] \mathcal{D}\mathbf{x}(t),\tag{3}$$

A very important step in calculation of this integral is to write the correct form of the Jacobian of transformation from stochastic to dynamical variables. It is the analytical factor related to this Jacobian of transformation that guarantees the optimal compensation of the noise-induced errors by providing a leading order contribution to the analytical expression for the posterior PDF. Another key is to introduce a convenient parameterizations of the vector field. We now briefly summarize the results of these calculations.

2.1 Parameterization of the vector field and likelihood construction

We assume that the data block of measurement $\mathcal{X} \equiv \{\mathbf{x}(t_k), k = 1, \dots, K\}$ comes from an uniform time grid:

$$\mathbf{x}_k \equiv \mathbf{x}(t_k) \quad (4)$$

$$t_k = t_0 + k h, \quad k = 1, \dots, K. \quad (5)$$

$$h = \frac{t_K - t_0}{K} \quad (6)$$

With this discretization, the system in (1) might be approximated as:

$$\begin{aligned} \dot{\mathbf{x}}_{k+1} &= \mathbf{x}_k + h \mathbf{f}(\mathbf{x}_k^* | \mathbf{c}) + \boldsymbol{\xi}_k \\ \langle \boldsymbol{\xi}_k \rangle &= 0, \quad \langle \boldsymbol{\xi}_s \boldsymbol{\xi}_q^T \rangle = \hat{\mathbf{D}} \delta_{sq} \end{aligned} \quad (7)$$

where use of the following definition was made:

$$\mathbf{x}_k^* = \frac{\mathbf{x}_{k+1} - \mathbf{x}_k}{2}$$

and the vector $\boldsymbol{\xi}_k = \int_{t_k}^{t_{k+1}} \boldsymbol{\xi}(s) ds$. on the uniform time lattice the discretized version of the logarithm of the likelihood takes the form:

$$\log \ell(\mathcal{X} | \mathcal{M}) = -\frac{K}{2} \ln \det \hat{\mathbf{D}} - \frac{h}{2} \sum_{k=0}^{K-1} \left[\text{tr} \left[\frac{\partial \mathbf{f}(\mathbf{x}_k^* | \mathbf{c})}{\partial \mathbf{x}} \right] + (\dot{\mathbf{x}}_k - \mathbf{f}(\mathbf{x}_k^* | \mathbf{c}))^T \hat{\mathbf{D}}^{-1} (\dot{\mathbf{x}}_k - \mathbf{f}(\mathbf{x}_k^* | \mathbf{c})) \right] + \mathcal{I}. \quad (8)$$

Here we have introduced the “velocity” $\dot{\mathbf{x}}_k$

$$\dot{\mathbf{x}}_k \equiv h^{-1}(\mathbf{x}_{k+1} - \mathbf{x}_k).$$

and \mathcal{I} is a constant factor which depends on neither \mathcal{M} nor \mathcal{X} . It should be noted that, in general, the likelihood in eq.(8) is such that the integral

$$\int \ell(\mathcal{X} | \mathcal{M}) p(\mathcal{M}) d\mathcal{M}$$

which appears in eq.(2) does not have a closed analytic solution. To overcome this problem authors of⁷ proposed the following parameterizations of the vector field:

$$\begin{aligned} \mathbf{f}(\mathbf{x} | \mathbf{c}) &\equiv \hat{\mathbf{U}}(\mathbf{x}) \mathbf{c} \\ \hat{\mathbf{U}}(\mathbf{x}) &\equiv \left[\begin{bmatrix} u_1(\mathbf{x}) & & \\ & \ddots & \\ & & u_1(\mathbf{x}) \end{bmatrix} \dots \begin{bmatrix} u_N(\mathbf{x}) & & \\ & \ddots & \\ & & u_N(\mathbf{x}) \end{bmatrix} \right] \end{aligned}$$

where $\{u_1(\mathbf{x}), \dots, u_N(\mathbf{x})\}$ is a set of suitable base functions. With this parameterization we assume that the function $\mathbf{f}(\mathbf{x})$, which is nonlinear in respect of \mathbf{x} , might be expressed as a weighted sum of nonlinear base functions. Parameters \mathbf{c} are the weights of this sum. Note that $\mathbf{f}(\mathbf{x}_k | \mathbf{c})$ might be highly non linear in \mathbf{x} , and we are only assuming linearity in respect of parameters.

In this way the likelihood function became a multivariate normal distribution in respect of \mathbf{c} . Thus, for a given $\hat{\mathbf{D}}$, taking the initial prior $p_{\text{pr}}(\mathbf{c})$ to be a multivariate normal distribution centered in some prior mean \mathbf{c}_{pr} and with prior covariance $\hat{\boldsymbol{\Sigma}}_{\text{pr}}$, the posterior $p_{\text{post}}(\mathbf{c})$ is also Gaussian. So the algorithm for finding both \mathbf{c} and $\hat{\mathbf{D}}$ might be a two step optimization process which performs iteratively the following two operations: (i) given \mathbf{c} , compute the best choice of $\hat{\mathbf{D}}$; (ii) given $\hat{\mathbf{D}}$, compute the best choice of \mathbf{c} . The details of the algorithm have been given elsewhere⁷ but a few comment will be made. The quantity to be maximized in respect of \mathbf{c} , and $\hat{\mathbf{D}}$ if the negative logarithm of the posterior is

$$S_{\mathcal{X}}(\mathbf{c}, \hat{\mathbf{D}}) = \frac{1}{2} \rho(\hat{\mathbf{D}}) - \mathbf{c}^T \mathbf{w}(\hat{\mathbf{D}}) + \frac{1}{2} \mathbf{c}^T \boldsymbol{\Xi}(\hat{\mathbf{D}}) \mathbf{c}. \quad (9)$$

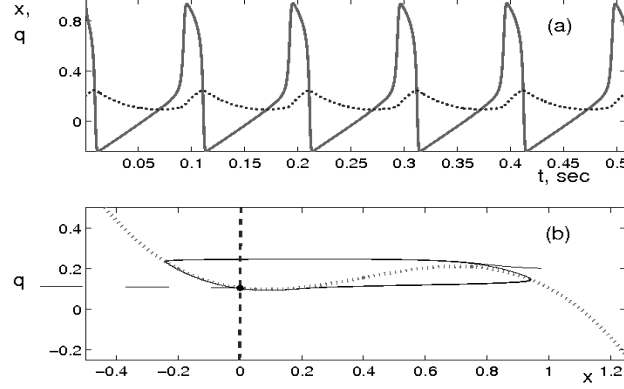


Figure 1. Synthetic data for one uncoupled FHN system. (a) $x(t)$ and $q(t)$, (b) limit cycles (solid line) and isoclines: $-x_1(x_1 - \theta_1)(x_1 - 1) - q_1 + \eta_1 = 0$ (dashed line); $\beta x_1 - \gamma_1 q_1 = 0$ (bold dashed line). Parameters are: $\gamma_1 = 0.0051051$, $\beta = 0.0051$, $\alpha_1 = 0.2$, $\eta_1 = 0.112$, and $d_1 = 1 \times 10^{-5}$.

where the following definitions has been used

$$\rho(\hat{\mathbf{D}}) \equiv h \sum_{k=0}^{K-1} \dot{\mathbf{x}}_k^T \hat{\mathbf{D}}^{-1} \dot{\mathbf{x}}_k + K \ln(\det \hat{\mathbf{D}}), \quad (10)$$

$$\mathbf{w}(\hat{\mathbf{D}}) \equiv \hat{\Sigma}_{\text{pr}}^{-1} \mathbf{c}_{\text{pr}} + h \sum_{k=0}^{K-1} \left[\hat{\mathbf{U}}^T(\mathbf{x}_k^*) \hat{\mathbf{D}}^{-1} \dot{\mathbf{x}}_k - \frac{1}{2} \sum_{l=1}^L \frac{\partial u_l(\mathbf{x}_k^*)}{\partial x_l} \right], \quad (11)$$

$$\Xi(\hat{\mathbf{D}}) \equiv \hat{\Sigma}_{\text{pr}}^{-1} + h \sum_{k=0}^{K-1} \hat{\mathbf{U}}^T(\mathbf{x}_k^*) \hat{\mathbf{D}}^{-1} \hat{\mathbf{U}}(\mathbf{x}_k^*). \quad (12)$$

where at each steps Ξ^{-1} is the new covariance matrix given $\hat{\mathbf{D}}$, and the next step parameter vector is $\mathbf{c}' = \Xi \mathbf{c}$, given $\hat{\mathbf{D}}$; while, given \mathbf{c} , the maximum of the negative logarithm of posterior in respect of $\hat{\mathbf{D}}$ is

$$\hat{\mathbf{D}} = \frac{h}{K} \sum_{k=0}^{K-1} [\dot{\mathbf{x}}_k - \hat{\mathbf{U}}_k \mathbf{c}]^T [\dot{\mathbf{x}}_k - \hat{\mathbf{U}}_k \mathbf{c}]. \quad (13)$$

3. THEORY OF DECOUPLING OF A NUMBER OF FITZHUGH-NAGUMO SYSTEMS MIXED BY MEASUREMENT MATRIX

In this section we will take into consideration the signal coming from a FitzHugh-Nagumo model. This is a widely used model for the neuron dynamic²⁻⁵ and, in particular, for nerve fibres⁶ and heart tissues.⁷⁻⁹ It consist of two coupled first order differential equation described by the variable x and q . x is the membrane potential, w is a recovery variable, slower than x .⁵ A set of multiple neurons can be written as:

$$\begin{aligned} \dot{x}_j &= -x_j(x_j - \alpha_j)(x_j - 1) - q_j + \eta_j + \xi_j, \\ \dot{q}_j &= -\beta q_j + \gamma_j x_j, \\ \langle \xi_j(t) \xi_i(t') \rangle &= d_{ij} \delta(t - t'), \quad i, j = 1 : L. \end{aligned} \quad (14)$$

j labels all the L neurons, and at the membrane potential a white Gaussian noise is present. Parameters η_j , α_j , γ_j are possibly different for each neuron. In particular η_j has a special importance since it models the input current at the membrane and it acts as a control parameter for the firing frequency, related to the information transfer in physiological systems. It is particularly important in the view of the physiological applications to decode this parameter.

Note that, in general, parameter β also has to be inferred. However, in what follows we will assume for the sake of simplicity that it is a known constant for each neuron.

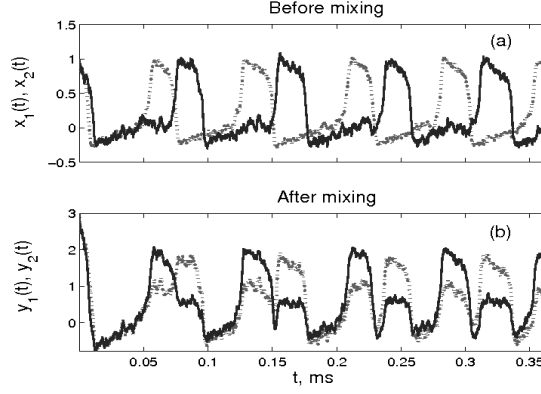


Figure 2. Model of measurements assumes that $x_1(t)$ and $x_2(t)$ are mixed with known matrix X as in (23). (a) $x_1(t)$ (black solid line) and $x_2(t)$ (grey dotted line), (b) $y_1(t)$ (black solid line) and $y_2(t)$ (grey dotted line). Parameter as in Table 4.1.

Note also that, in practice, the signal from different neurons are mixed by a measurement matrix X , so it would be useful to have a technique to decode these parameter from a corresponding readout variable related to the membrane potentials via the following expression:

$$y_i = X_{ij} x_j \quad (15)$$

Where y_i represents the output of our signal, i.e. see Fig. (2). With this approach we have to find the basis functions $u(y_1, \dots, y_L)$ such that the systems

$$\dot{y}_i = \sum_j a_{ij} u_j(y_1, \dots, y_L)$$

represents the dynamics of measured variables corresponding to the underlying dynamics of actual variables described by eq.(14) and measurement model given by (3).

First q has to be reconstructed, since it is not read directly. We solve equation eq.(14), that is:

$$q_j(t) = \gamma \int_0^t d\tau \exp(-\beta(t-\tau)) x_j(\tau) + \exp(-\beta t) q_j(0)$$

Here $q_j(0)$ is a set of initial coordinates that needs to be inferred along with the rest of parameters. We plug this equation into eq.(14) and obtain:

$$\dot{x}_j(t) = -x_j(x_j - \alpha_j)(x_j - 1) + \eta_j + d_j - \gamma \int_0^t \exp(-\beta(t-\tau)) x_j(\tau) d\tau + \exp(-\beta t) q_j(0) . \quad (16)$$

The function $\int_0^t \exp(-\beta(t-\tau)) x_j(\tau) d\tau$ is defined over the time grid $\{t_0, t_1, \dots, t_N\}$. After some algebraic manipulations, using trapezoidal rule for the evaluation of the integral, it can be shows that

$$\int_0^{t_k} \exp(-\beta(t-\tau)) x_j(\tau) d\tau = h \sum_{n=0}^N x_j(t_n) \exp(\beta(t_n - t_N)) - \frac{h}{2} (x_j(t_0) \exp(\beta(t_0 - t_N)) + x_j(t_N)) .$$

Substituting eq.(16) into eq.(15) we obtain finally the required system of equations in the form:

$$\begin{aligned} \dot{y}_j = & \sum_{i=1}^N A_{ji} y_i + \sum_{k1, k2=1}^N B_{jk1k2} y_{k1} y_{k2} + \sum_{k1, k2, k3=1}^N C_{jk1k2k3} y_{k1} y_{k2} y_{k3} + \\ & - \int_{t_0}^t d\tau \sum_{i,k=1}^N \exp[-\beta(t-\tau)] \Gamma_{ik}^j y_k(\tau) - \sum_{i=1}^N \exp[-\beta t] \tilde{q}_j + \tilde{\eta}_j + \sum_{i=1}^N D_{ji} \xi_i(t). \end{aligned} \quad (17)$$

Here, use was made of the following definitions:

$$\begin{aligned}
A_{j\ k} &= \sum_{i=1}^N X_{j\ i} \alpha_i (X^{-1})_{i\ k}, \\
B_{j\ k_1\ k_2} &= \sum_{i, k_1, k_2=1}^N X_{j\ i} (1 + \alpha_i) (X^{-1})_{i\ k_1} (X^{-1})_{i\ k_2} \\
C_{j\ k_1\ k_2\ k_3} &= \sum_{i, k_1, k_2, k_3=1}^N X_{j\ i} (X^{-1})_{i\ k_1} (X^{-1})_{i\ k_2} (X^{-1})_{i\ k_3}, \\
\Gamma_{i\ k}^j &= X_{j\ i} \gamma_i (X^{-1})_{i\ k}.
\end{aligned} \tag{18}$$

The modified noise intensities D_j for an auxiliary system (3) are expressed in terms of d_j , and the modified initial condition for \tilde{q}_i are expressed in terms of q_j as follows:

$$D_j = \sum_{i=1}^N X_{j\ i} d_i \quad \tilde{q}_j = \sum_{i=1}^N X_{j\ i} q_i(0). \tag{19}$$

The parameters $\tilde{\eta}_j$ that appear in eq.(17) are related to the original model parameters η_j , as follows:

$$\tilde{\eta}_j = \sum_{i=1}^N X_{j\ i} \eta_i. \tag{20}$$

From eq.(17), we can see that if the system's dimension L increases, and if we cannot make any assumption on the form of the coupling, the increase in the number of parameters is proportional to L^3 . In aprticular, when $L = 2$, the following base functions have to be used to infer matrix elements of $A, B, C, \Gamma, D, \tilde{\eta}$:

$$\begin{aligned}
u_1 &\equiv 1 & u_2 &\equiv y_1 & u_3 &\equiv y_2 & u_4 &\equiv y_1^2 & u_5 &\equiv y_2^2 \\
u_6 &\equiv y_1 y_2 & u_7 &\equiv y_1^3 & u_8 &\equiv y_1^2 y_2 & u_9 &\equiv y_1 y_2^2 & u_{10} &\equiv y_2^3 \\
u_{11} &\equiv h \sum_{l=0}^k y_1(l) \exp(\beta(t_l - t_k)) - \frac{h}{2} (y_1(0) \exp(\beta(t_0 - t_k)) + y_1(k)), \\
u_{12} &\equiv h \sum_{l=0}^k y_2(l) \exp(\beta(t_l - t_k)) - \frac{h}{2} (y_2(0) \exp(\beta(t_0 - t_k)) + y_2(k)) \\
u_{13} &\equiv -\exp(\beta(t_0 - t_k)).
\end{aligned} \tag{21}$$

Thus the total number of parameters is 26 and the vector \mathbf{c} as a function of the parameters' matrix elements takes the form

$$\begin{aligned}
\mathbf{c} &= [\tilde{\eta}_1, \tilde{\eta}_2, -A_{11}, -A_{21}, -A_{12}, -A_{22}, \dots \\
&\quad B_{111}, B_{211}, (B_{112} + B_{121}), \dots \\
&\quad (B_{212} + B_{221}), B_{122}, B_{222}, -C_{1111}, -C_{2111}, \dots \\
&\quad -(C_{1112} + C_{1121} + C_{1211}), -(C_{2112} + C_{2121} + C_{2211}), \dots \\
&\quad -(C_{1122} + C_{1221} + C_{1212}), -(C_{2122} + C_{2221} + C_{2212}), \dots \\
&\quad -C_{1222}, -C_{2222}, \gamma_1, \gamma_2, \tilde{q}_1, \tilde{q}_2].
\end{aligned} \tag{22}$$

4. INFERENCE OF THE CONTROL PARAMETERS FOR TWO COUPLED FHN SYSTEMS

In this section the inference algorithm described above will be tested using synthetically generated signal for the system of two FHN oscillators. First, we will describe how the signal is generated and how the stream of data

is passed on to the inferring machinery; then tests of convergence will be carried out in the case of a stationary dynamic; finally a demonstration and consideration of how the technique performs in the case of a non-stationary dynamics will be presented.

4.1 Synthetic data and preliminary test

To test our inferential algorithm we generate 2 signals from a 2-dimensional FHN system as in eq.(14), keeping all parameters fixed. To generate the stream of data, the stochastic differential equation has been integrated accordingly to the Heun scheme.¹⁰ In Fig. (2) (a) a sample of data generated with coefficient as in Table (4.1) is shown. The signals y_i obtained multiplying the mixing matrix with the signals x_i is the input for the algorithm.

$\alpha_1 = 0.2$	$\eta = 0.15$
$\alpha_2 = 0.2$	$\eta = 0.15$
$\beta_1 = 0.0051$	$\gamma_1 = 0.0051051$
$\beta_2 = 0.0051$	$\gamma_2 = 0.0051051$
$d_{11} = 0.004$	$d_{12} = 0.001$
$d_{22} = 0.004$	$d_{21} = 0.001$

Table 1. Values of coefficients for the generated stationary signal

The measurement matrix has been taken in the following form:

$$X = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad (23)$$

In this way we generate contiguously multiple blocks of points. Values of the control parameters η_1 and η_2 change step-wise at random from block to block and remain constant within each block. Other parameters of the system remain constant all the time. The inference is performed through the following steps:

1. Use a non-informative prior for parameters with an infinitely large variance.
2. Compute the first block $p_{\text{post}}^{\text{1st block}}(\mathcal{M})$.
3. Reset the prior to an infinitely large multivariate gaussian distribution for each block of data.

In Figs.(3)(a) and (b) inferred values of the parameter η_i are compared with their actual values for five time steps. The sampling rate was approximately 50 kHz, i.e. we use 20000 points to infer values of the model parameters on one time step. As it is shown in the figure, the time interval between steps is approximately 5-10 periods of firing of action potential. The measured trajectories $y_1(t)$ and $y_2(t)$ are shown in the Fig. (3) (b) and (e). In the Fig. (3) (c) and (f) we compare the reconstructed dynamic variable q_i for each oscillator. In order to obtain the unmeasured coordinate $q_i(t)$ the second line of eq.(14) has been integrated numerically using inferred coefficients, and using the boundary condition $q_i(0)$ that can be derived from eq.(19):

$$q_i(0) = \sum_{j=1}^N X_{ij}^{-1} \tilde{q}_j \quad . \quad (24)$$

4.2 Test of convergence

It is also advisable to check how the convergence depends on the total time of measurements of the time-series data. To this end we generate synthetic time-series data keeping all coefficients constant and express the total measurement time as the number of blocks. A ensemble of different sample signals (number of *runs*) is generated for each value of the total measurement time and inference procedure is carried out independently for each run. Statistical properties from the ensemble of inferred parameters computed in this way provide an example of the convergence of coefficients c_i as shown in Fig. (4). The number of runs to obtain the averaged convergence was 3000 for each size of the block of data. It can be seen from the figure that the time of convergence is larger

for the coefficients corresponding to higher powers of the dynamical variables in the equations. Convergence of the diffusion coefficients is shown in Fig. (5). As it can be seen that the inferred parameters for diffusion matrix $\hat{\mathbf{D}}$ are on average overestimated compared to the real ones. This behavior is rather simply explained: the maximum of the likelihood in respect of $\hat{\mathbf{D}}$ is given by eq.(13). Nevertheless it holds for real coefficients \mathbf{c} . In our case the vector \mathbf{c} is not exact, but it represents the maximum of the likelihood in respect of vector \mathbf{c} , thus it is a statistical quantity. Thus, the slight bias in the estimation of $\hat{\mathbf{D}}$ reflects the propagation of the uncertainty in the inferred \mathbf{c} . To avoid this bias one has to repeat calculations of $\hat{\mathbf{D}}$ and \mathbf{c} in a loop until convergence is reached as will be explained in details elsewhere.

The coefficients presented in Figs. (4) and (5) are the weight factors for the base functions of eq.(21). To reconstruct the original coefficients, and to check the quality of our estimation, we recover coefficients (14) and the elements of the diffusion matrix d_i using eqs.(20) that are reduced in our case to the following expressions:

$$\mathcal{A} = X^{-1}AX, \quad (25)$$

$$\Gamma = X^1 \tilde{\Gamma} X, \quad (26)$$

$$D = X^{-1} \tilde{D} (X^T)^{-1}, \quad (27)$$

$$\eta = X^{-1} \tilde{\eta}. \quad (28)$$

Here \mathcal{A} is a diagonal matrix with diagonal elements equal to α_1 and α_2 , Γ is a diagonal matrix with diagonal elements equal to γ_1 and γ_2 . Since in our test $\alpha_1 = \alpha_2$, $\gamma_1 = \gamma_2$ and $\eta_1 = \eta_2$ we present the results of recovering of the original coefficients only for the first equation. The results of the inference of the original equations obtained from the inference of the mixed equations with the help of relations (28) are shown in Fig. (6). For the block of data containing 8000 points, results of the inference are summarized in Table 2.

4.3 Parameters η

If we devote our attention only to the inference of parameters η_1 and η_2 , namely if prior knowledge of all the other parameters is given, clearly we need much less information (sample data) to achieve good estimations of η_i . As has been done formerly, η_1 and η_2 change step-wise at random and remain constant between steps. However, the time interval between steps is now 20 times smaller and is less than one period of firing of an action potential.

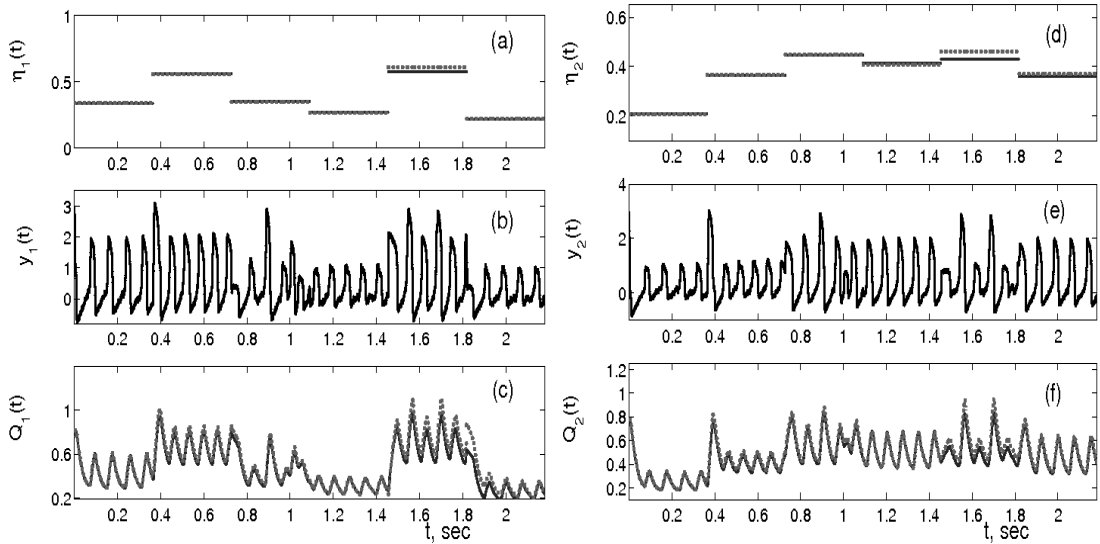


Figure 3. Inference of the model parameters of two uncoupled FHN systems mixed by the measurement matrix with step-wise changes of η_1 and η_2 . (a) Actual values of η_1 are shown by solid lines in comparison with inferred values shown by dotted lines. (b) Measured mixed values of coordinates $y_1(t)$ (solid line). (c) Mixed value of the generated coordinate $q_1(t)$ is shown in comparison with its inferred value integrated from \tilde{q}_i . Figures (d), (e), and (f) show the corresponding result for the second system. The values of other parameters were $\alpha_1 = \alpha_2 = 0.2$, $\beta = 0.0051051$, $\gamma_1 = \gamma_2 = 0.0051$.

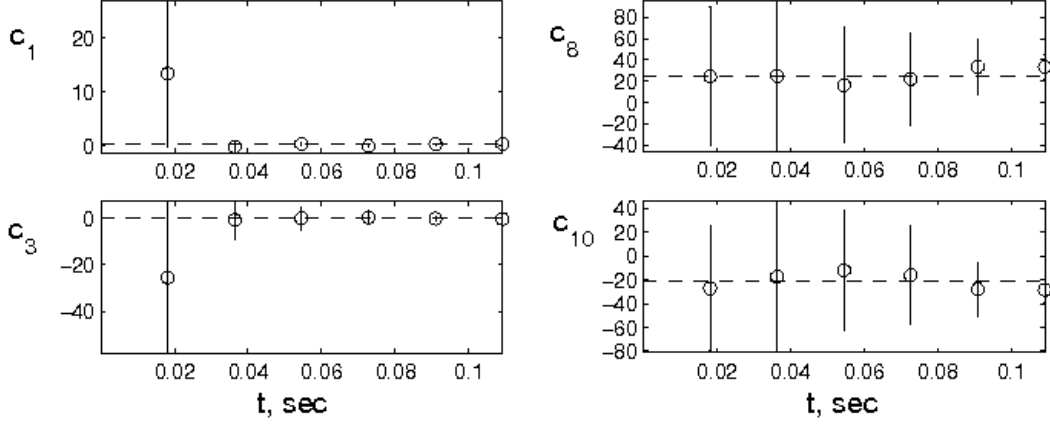


Figure 4. Convergence of some of the model parameters c_i as a function of the length of the block of data. The sampling rate is approximately 50 kHz. First point corresponds to the 1000 data points in one block. For each next point the number of data points was increased by 1000. Vertical lines shows standard deviation of the inferred values of the model coefficients. The horizontal dashed lines shows actual values of the model parameters. The values of other parameters were $\alpha_1 = \alpha_2 = 0.2$, $\beta = 0.0051051$, $\gamma_1 = \gamma_2 = 0.0051$.

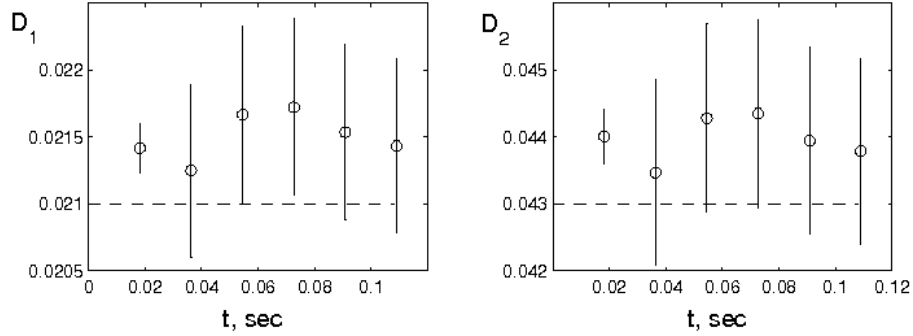


Figure 5. Convergence of the diffusion coefficients as a function of the length of the block of data. Parameters are the same as in Fig. (4).

Other parameters of the system remain constant all the time. At each step we infer only parameters η_1 and η_2 of the model assuming their initial values to be zero and their initial dispersion to be infinity. In Fig. (7) inferred values of the parameter η_i are compared with their actual values for five time steps. The sampling rate was approximately 50 kHz, i.e. we use 1000 points to infer values of the model parameters in each time step. The measured trajectories $y_1(t)$ and $y_2(t)$ are shown in the Fig. (7) (b) and (e). The values of the real and inferred coordinates $q_i(t)$ are compared in Fig. (7) (c) and (f).

The results obtained so far show a remarkably close correspondence between the real and inferred values, but one can argue that this correspondence is due to the fact that real parameters η_i change stepwise but remain constant during each block of inference. Therefore, in the following we will investigate the possibility of inference in case of continuous smoothly-varying parameters η_1 and η_2 that have both periodic and random components. Although the algorithm is the same, the data set now consists of a series of blocks 10000 points long each; i.e. we have a few periods of firing per one block of inference. We also assume that all other parameters of the system are known. The results of inference are shown in the Fig. (8). The results are very similar to those shown in the Fig. (7) except that we now infer the average value of η_i per block. This result suggests that we can obtain a better approximation of the actual values of η by shifting the initial time of each block and assuming that the inferred value of η_i corresponds to the time at the middle point of the block.

It is clear that thanks to this “block by block” approach the quality of inference for smoothly-varying parameters seems to fulfill the researcher’s expectation; thus, to complete the scenario of the inference of multiple

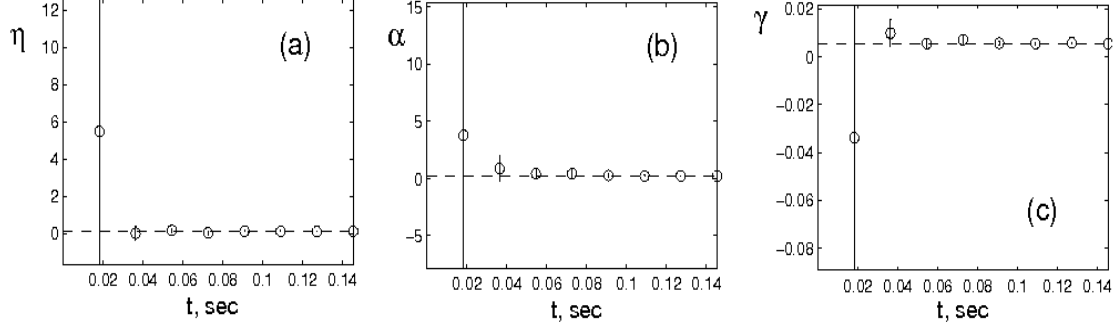


Figure 6. Convergence of the original coefficients (a) η_1 , (b) α_1 , (c) γ_1 , and (d) d_1 as functions of the time if inference was conducted with a sampling rate of 55 kHz. Parameters are the same as in Fig. (4).

η_1	α_1	γ_1	d_1
0.112	0.2	0.0051	0.001
0.1245	0.2161	0.0055	0.0010
0.0097	0.0202	0.0003	0.0003

Table 2. Values of the original coefficients inferred using 8000 points measured with sampling rate 55 kHz obtained from mixed measurements with the help of eqs. (28). The actual values (top row) are compared with the inferred values (middle row), standard deviations are given in the bottom row.

parameters, we present the performances of the algorithm in case when all parameters need to be inferred.

Hence, in the next and last part no assumptions are made, nor prior knowledge of others parameters taken as given, and η_1 and η_2 are varying with periodic and random components. The results for a block of data containing 8000 points results of inference are summarized in Table 3. At each step of the algorithm we use the

c_1	c_2	c_3	c_4	c_5	D_{11}
0.1106	0.0607	-0.2000	0	0	0.0009
0.2084	0.2074	-0.2130	-0.0062	0.0130	0.0010
c_7	c_8	c_9	c_{10}	c_{11}	D_{12}
1.2000	0.8000	-1.6000	-1.6000	0.8000	0.0006
1.2174	0.9871	-1.5712	-2.0007	-1.0123	0.0006
c_{18}	c_{19}	c_{20}	c_{21}	c_{22}	D_{22}
0.6667	-0.2222	-0.5556	0.0051	0.7500	0.0009
1.5529	-0.3399	-0.7994	0.0049	0.7255	0.0010

Table 3. Values of some model coefficients inferred at the first step. The actual values (top row) are compared with the inferred values (bottom row).

coefficients of the model inferred at the previous step to infer η_i parameters. We use a sampling rate of 55 kHz and 15000 points per block of inference. The results are shown in Fig. (9). They are very similar to those shown of Fig. (8) even though the values of the model coefficients are now known only approximately, thus confirming on the efficacy of the technique for parameter-tracking purposes.

5. CONCLUSION

In this paper we have presented a new application of the Bayesian dynamical inference technique. To the previously proposed method¹ additional features have been added, namely we have taken into account the variability of parameters of the model that has to be inferred. Here (a) driving parameters that are varying in time have been inferred in a dynamical system of coupled FHN oscillators with additive noise, and (b) the general framework of application the Bayesian algorithm to a multiple FHN system has been presented.

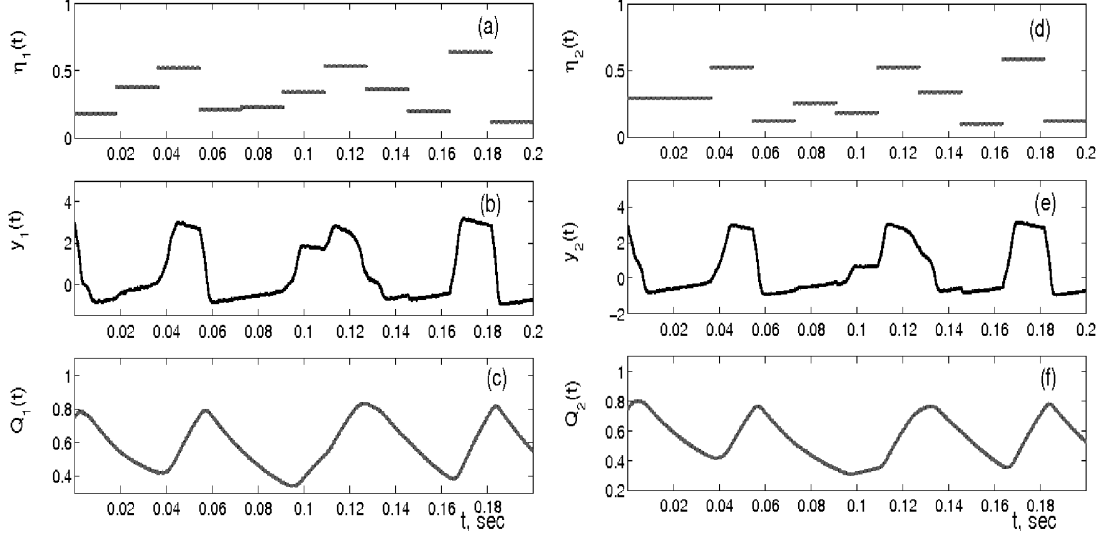


Figure 7. Inference of η_1 and η_2 for two uncoupled FHN systems mixed by the measurement matrix. (a) The actual values of η_1 , are shown by solid lines, are almost indistinguishable from the inferred values, shown by dotted lines. (b) Measured mixed values of coordinates $y_1(t)$ (solid line). (c) Values of the coordinate $q_1(t)$ are shown in comparison with its inferred value. Figures (d), (e), and (f) show the same results for the second system. The values of other parameters where $\alpha_1 = \alpha_2 = 0.2$, $\beta = 0.0051051$, $\gamma_1 = \gamma_2 = 0.0051$.

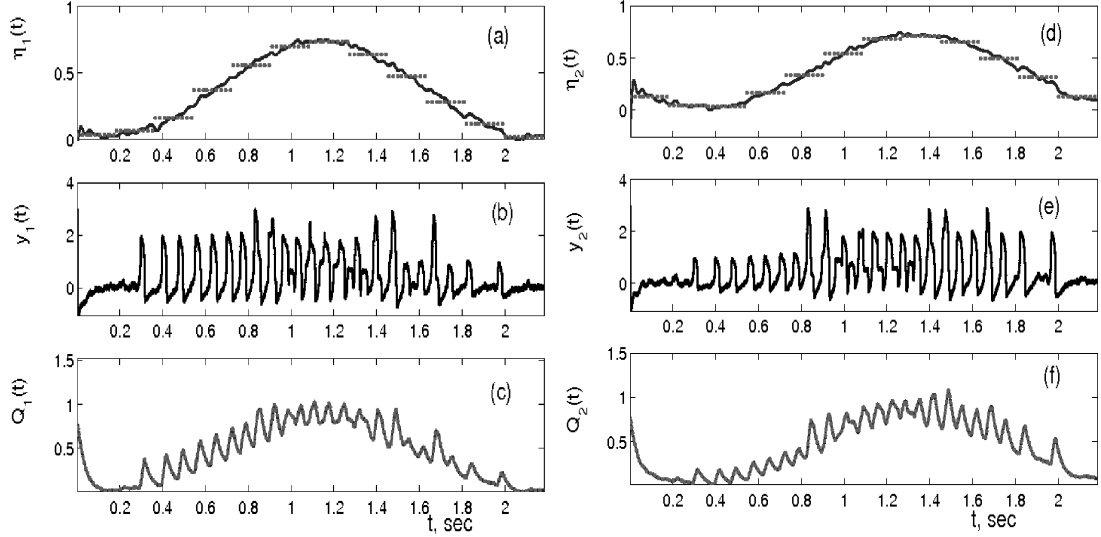


Figure 8. Inference of smoothly-varying η_1 and η_2 that have both periodic and random components for two uncoupled FHN systems. (a) Actual values of η_1 are shown by black solid lines in comparison with inferred values shown by gray dotted lines. (b) Measured mixed values of coordinates $y_1(t)$ (solid line). (c) Values of the coordinate $q_1(t)$ (solid line) is almost indistinguishable from the inferred value (gray dotted line). Figures (d), (e), and (f) show the same results for the second system. The values of other parameters where $\alpha_1 = \alpha_2 = 0.2$, $\beta = 0.0051051$, $\gamma_1 = \gamma_2 = 0.0051$.

As readout variables we considered a vector formed from a linear combination of the system dynamical variables mixed through a measurement matrix. The multiple FHN system here is therefore taken in general form. This allows us to demonstrate that it is possible to recover the real system parameters and the diffusion coefficients.

As a result we have shown that new algorithm is advantageously fast in detecting dynamic controlling parameters, and it appears to be a desirable and flexible instrument in analysing the data, because it maximises the use of

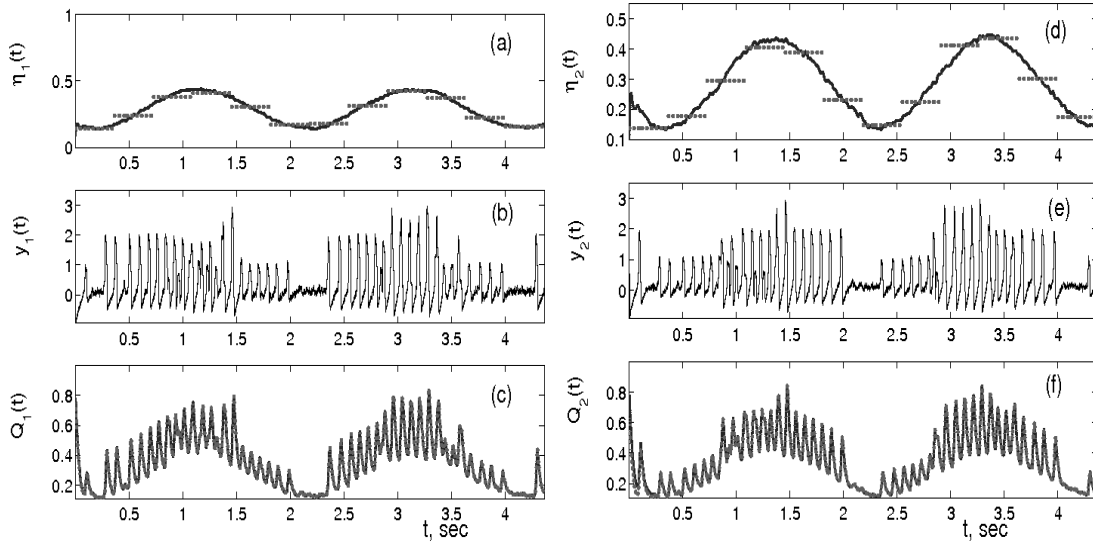


Figure 9. Inference of smoothly-varying η_1 and η_2 that have both periodic and random components for two uncoupled FHN systems using model coefficients inferred at the first step. (a) Actual values of η_1 are shown by black solid lines in comparison with inferred values shown by gray dotted lines. (b) Measured mixed values of coordinates $y_1(t)$ (solid line). (c) Values of the coordinate $q_1(t)$ are almost indistinguishable from the inferred value (gray dotted line). Figures (d), (e), and (f) show the same results for the second system. The values of other parameters where $\alpha_1 = \alpha_2 = 0.2$, $\beta = 0.0051051$, $\gamma_1 = \gamma_2 = 0.0051$.

the available information. Indeed, we have demonstrated it by performing tests where stepwise changes of the dynamics were read under several hypotheses about the factual external knowledge.

Due to these properties our inference algorithm is also able to follow continuous changes of a system's parameters. In fact it makes it possible to refresh the inferred system's state instantaneously (providing opportunity for the model to reflect parameter changes on a continuous time basis).

REFERENCES

1. V. N. Smelyanskiy, D. G. Luchinsky, D. A. Timucin, and A. Bandrivskyy, "Reconstruction of stochastic nonlinear dynamical models from trajectory measurements," *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* **72**(2), p. 026202, 2005.
2. R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical Journal* **1**(6), pp. 445–466, 1961.
3. J. Nagumo, S. Animoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," *Proc. Inst. Radio Engineers* **50**, pp. 2061–2070, 1962.
4. A. T. Winfree, *The Geometry of Biological Time*, Springer-Verlag, New York, 1980.
5. E. Izhikevich, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting.*, MIT Press, 2005. preprint.
6. J. Keener and J. Sneyd, *Mathematical Physiology*, Springer-Verlag, New York, 1998.
7. E. N. Best, "Null space in the hodgkin-huxley equations - a critical test," *Biophysical Journal* **27**(1), pp. 87–104, 1979.
8. J. Rogers and A. McCulloch, "A collocation-galerkin finite element model of cardiac action potential propagation," *IEEE Trans. Biomed. Eng.* **41**, pp. 743–757, 1994.
9. R. R. Aliev and A. V. Panfilov, "Modeling of heart excitation patterns caused by a local inhomogeneity," *Journal of Theoretical Biology* **181**(1), pp. 33–40, 1996.
10. R. Mannella, "Integration of stochastic differential equations on a computer," *International Journal of Modern Physics C* **13**(9), pp. 1177–1195, 2002.